

CLAIMS

1. Apparatus for creating and maintaining a user profile for a user for improving database searching by the user, said apparatus comprising:
- 5 means for accessing a predetermined set of documents containing a plurality of keywords during a learning phase;
- analysing means arranged to analyse said documents and to identify, according to predetermined rules, groups of related keywords therein;
- attribute assigning means arranged to assign attributes indicative of
- 10 relatedness to said groups of keywords; and
- user profile storing means arranged to store said relatedness attributes as a user profile;
- said apparatus further comprising:
- document updating means arranged to update the set of documents by
- 15 adding documents to or subtracting documents from the set during an updating phase;
- identifying means arranged to analyse the updated set of documents and to identify existing and additional groups of related keywords therein, according to predetermined rules;
- 20 means arranged to assign attributes indicative of relatedness to said additional groups of keywords;
- relatedness attribute updating means for updating the relatedness attributes of said existing groups of keywords; and
- user profile updating means arranged to update the user profile in accordance
- 25 with the relatedness attributes of said existing and additional groups of keywords.
2. Apparatus according to claim 1, wherein the predetermined set of documents is a set of documents expected to reflect the interests of a specific user.
- 30 3. Apparatus according to claim 1 or 2, wherein the predetermined set of documents is a set of documents derived from a set of documents previously viewed by a specific user.

4. Apparatus according to claim 1, 2 or 3, wherein the analysing means comprises means for identifying groups containing pairs of related keywords.
5. Apparatus according to claim 1, 2, 3 or 4, wherein the analysing means
5 comprises means for identifying related keywords from the set of documents by means of a self-organising map algorithm.
6. Apparatus according to any of claims 1 to 5, wherein the attribute assigning means comprises importance value assigning means for assigning importance values
10 indicating the statistical significance of related keywords in the set of documents.
7. Apparatus according to any of claims 1 to 6, wherein the attribute assigning means comprises means for assigning life-span values indicating the expected remaining period of time of relatedness between keywords in the set of documents.
15
8. Apparatus according to claim 7, wherein said relatedness attribute updating means comprises means for systematically decreasing the life-span values over time.
9. Apparatus according to any of claims 1 to 8, wherein the document updating
20 means is arranged to update the set of documents in response to user input.
10. Apparatus according to claim 9, wherein the document updating means is arranged to add new documents to the set of documents in the event of user input confirming that said new documents are of interest to the user.
25
11. Apparatus according to any of claims 1 to 10, wherein the user profile storing means is arranged to store said relatedness attributes in the form of fuzzy sets.
- 30 12. A method for creating and maintaining a user profile for a user for improving database searching by the user, said method comprising a learning phase and an updating phase, wherein said learning phase comprises the steps of:

accessing a predetermined set of documents containing a plurality of keywords;

analysing said documents and identifying, according to predetermined rules, groups of related keywords therein;

5 assigning attributes indicative of relatedness to said groups of keywords; and
storing said relatedness attributes as a user profile;

and wherein said updating phase comprises the steps of:

updating the set of documents by adding documents to or subtracting documents from the set;

10 analysing the updated set of documents and identifying existing and additional groups of related keywords therein, according to predetermined rules;

assigning attributes indicative of relatedness to said additional groups of keywords;

updating the relatedness attributes of said existing groups of keywords; and

15 updating the user profile in accordance with the relatedness attributes of said existing and additional groups of keywords.

13. A method according to claim 12, wherein groups containing pairs of related keywords are identified.

20

14. A method according to claim 12 or 13, wherein related keywords are identified from the set of documents by means of a self-organising map algorithm.

15. A method according to claim 12, 13 or 14, wherein the step of assigning
25 attributes comprises assigning importance values indicating the statistical significance of related keywords in the set of documents.

16. A method according to any of claims 12 to 15, wherein the step of assigning
30 attributes comprises assigning life-span values indicating the expected remaining period of time of relatedness between keywords in the set of documents.

17. A method according to claim 16, wherein the step of updating the relatedness attributes comprises a step of systematically decreasing the life-span values over time.

5 18. A method according to any of claims 12 to 17, wherein the step of updating the set of documents comprises updating the set of documents in response to user input.

19. A method according to claim 18, wherein the step of updating the set of
10 documents comprises adding new documents to the set of documents in the event of user input confirming that said new documents are of interest to the user.

20. A method according to any of claims 12 to 19, further comprising a step of
15 updating the set of documents on the basis of documents viewed by the user following receipt of a response from a search engine to a search query.

21. A method according to any of claims 12 to 20, wherein said relatedness attributes are stored in the form of fuzzy sets.

20 22. Apparatus for improving database searching, comprising:

user profile means, having access to a predetermined set of documents, arranged to provide data indicative of relatedness criteria between keywords from the set of documents;

25 means for receiving a search query comprising one or more search keywords from a user;

means arranged to access said user profile means and to identify therefrom, for the or each search keyword, potentially-related keywords according to predetermined criteria;

30 means arranged to provide said potentially-related keywords to the user;

means for receiving information from the user confirming that any potentially-related keywords are considered to be related keywords;

means arranged to incorporate such potentially-related keywords as keywords in an improved search query in the event that they are confirmed by the user to be related keywords; and

means for submitting the improved search query to a search engine.

5

23. Apparatus according to claim 22, wherein the predetermined set of documents is a set of documents expected to reflect the interests of a specific user.

24. Apparatus according to claim 22 or 23, wherein the predetermined set of documents is a set of documents derived from a set of documents previously viewed by the user.

25. Apparatus according to claim 22, 23 or 24, wherein the user profile means comprises means for identifying related keywords from the set of documents by means of a self-organising map algorithm.

26. Apparatus according to claim 22, 23, 24 or 25, wherein the user profile means comprises importance value deriving means for deriving importance values indicating the statistical significance of related keywords in the set of documents.

20

27. Apparatus according to any of claims 22 to 26, wherein the user profile means comprises means for assigning life-span values indicating an expected period of time of relatedness between keywords in the set of documents.

28. Apparatus according to any of claims 22 to 27, wherein the user profile means is arranged to provide said data indicative of relatedness criteria in the form of fuzzy sets.

29. Apparatus according to any of claims 22 to 28, further comprising means for updating the set of documents on the basis of documents viewed by the user following receipt of a response from a search engine to a search to a search query.

30

30. Apparatus according to any of claims 22 to 29, wherein the user profile means further comprises means for updating the data indicative of relatedness criteria on the basis of information received from the user.

5 31. A method for improving database searching, comprising the steps of:
receiving a search query comprising one or more search keywords from a user;

accessing a user profile means arranged to provide data indicative of relatedness criteria between keywords from a set of documents, and identifying from
10 said user profile means, for the or each search keyword, potentially-related keywords according to predetermined criteria;

providing said potentially-related keywords to the user;

receiving information from the user confirming that any potentially-related keywords are considered to be related keywords;

15 in the event that any potentially-related keywords are confirmed by the user to be related keywords, incorporating such potentially-related keywords as keywords in an improved search query; and

submitting the improved search query to a search engine.

20 32. A method according to claim 31, wherein the user profile means is arranged to identify said data indicative of relatedness criteria by means of a self-organising map algorithm.

33. A method according to claim 31 or 32, wherein the user profile means is
25 arranged to provide importance values indicating the statistical significance of related keywords in the set of documents.

34. A method according to claim 31, 32 or 33, wherein the user profile means is
30 arranged to provide life-span values indicating an expected period of time of relatedness between keywords in the set of documents.

35. A method according to any of claims 31 to 34, wherein the user profile means is arranged to provide said data indicative of relatedness criteria in the form of fuzzy sets.

5 36. A method according to any of claims 31 to 35, further comprising the step of updating the set of documents on the basis of documents viewed by the user following receipt of a response from a search engine to a search to a search query.

37. A method according to any of claims 31 to 36, further comprising the step
10 of updating the data indicative of relatedness criteria on the basis of information received from the user.